

A TCP AWARE LOCAL RETRANSMISSIONER SCHEME FOR UNRELIABLE TRANSMISSION NETWORK

BACKGROUND OF THE INVENTION:

Field of the Invention:

The present invention relates to a retransmission scheme on an unreliable network. More precisely, it concerns one kind of a TCP Aware Local Retransmission scheme for Transport Control Protocol (TCP). The scheme provides reliable retransmission on an unreliable transmission link and avoids effectively the disturbance between different layers of protocol, so it improves TCP performance of the unreliable network.

With the rapid development of the Internet, it is very possible it may be the united platform of multimedia networks in the future. Therefore, it is very important to access the Internet with various network technologies, among them accessing through a mobile communication system to the Internet. ITU is now making great efforts to promote international standards of the third generation mobile communication system (IMT2000), which is expected to support mobile computing and nomadic access of the Internet. It could be said, with the development of network technology and business, there are many varieties of the lower layers of the Internet. It was unpredictable for the Internet creators that the simple assumptions of the network characteristic at the creation beginning stage cause many of its kernel protocols are not adequate to the varieties of mixed network in the future, especially for the TCP layer.

Filed by Express Mail
(Receipt No. EC02102848)
on August 10, 2001
pursuant to 37 C.F.R. 1.10.
by B. B. B. B.

TCP is originally designed for wire link networks and has been widely used in the Internet. It provides reliable transport service for applications, such as WWW browsing, remote login (Telnet) and file transfer (FTP) etc. But many of the basic technologies, which the protocol is based on, are unreliable from the present point of view. For
5 example, all of the packet losses are regarded as the result of congestion.

Generally speaking, TCP performance is guaranteed by the technical schemes: the congestion control and error recovery. The congestion control is based on the sliding window scheme, which has been improved repeatedly and is nearly perfect. The basic idea of error recovery is the feedback of acknowledgement, with the transfer of limited
10 feedback information, i.e. the ACK (acknowledge) and NACK (non-acknowledge), to decide whether to retransmit the data packet. But it cannot differentiate the congestion loss and non-congestion loss by itself. This means in spite of loss being congestion or non-congestion, the TCP source inflates the window size automatically when ACK is received, and shrinks automatically when NACK received. In this way the window size is
15 adjusted to the network volume (throughput) in order to decrease the possibility of congestion. In practice of transmission processes especially in an unreliable network such as a wireless link, because the error rate is higher at the wireless environment and non-congestion loss is severe at a bad channel, this is an incorrect trigger of congestion control. This incorrect control causes unnecessary adjustment of window size, it
20 decreases availability of network resources. At the same time as the window is severely disturbed, TCP performance is worse. Therefore, TCP is facing many challenges in an unreliable network such as a wireless network.

On the other hand, the TCP recovery ability of losing packet is limited. There are many improving proposals for increasing the packet retransmission rate and efficiency, such as TCP-New Reno (Reference [1] J. C. Hoe, Improving the Start-Up Behavior of a Congestion Control Scheme for TCP. In proceeding ACM SIGCOMM 96, August 1996) and TCP-SACK (Reference [2] Fall and S. Floyd, Simulation-based Comparisons of Tahoe, Reno and Sack TCP, Computer Communication Review, July 1996). All of these proposals do not solve the difficult problem that is to identify the reason of losses. If using the Explicit Loss Notification (ELN) bit to solve this problem (Reference [3] Harl, Balakrishnan, and Randy H. Katz, Explicit Loss Notification and Wireless Web Performance, Globecom 98, Sydney, Australia, November 1998), it needs to update the kernel protocol of Internet, so it is infeasible.

In summary, a mixed network, especially with wireless networks, the urgent problem to be solved is to differentiate the congestion loss packet and non-congestion loss packet, and to solve the incorrect window size adjustment when non-congestion loss packet is detected.

In order to solve these problems, some local solutions have been proposed, such as based on link-layer AIRMAIL (Reference [4] E. Ayanoglu, S. Paul, T. F. Laporta, K. K. Sabnaai and R. D. Gitlin, AIRMAIL: A Link-Layer protocol for Wireless Network. ACM ACM/Baltzer Wireless Networks Journal, 1: 47-60, February 1995) and other ARQ/FEC (Automatic retransmission request. Forward error correction). With retransmission of the lost packet, they improve TCP performance, but with these pure local solutions, a large amount of overhead is added and decreases the efficiency. Reference [5] (H. Balakrishnan, S. Seshan and R. H. Katz, Improving Reliable Transport

and Handoff Performance in Cellular Wireless Networks, ACM Wireless Networks, December 1995) proposes the awareness concept to solve this problem. This concept applies TCP/ACK (Transmission Control Protocol/Acknowledgement) packet in the link layer to provide the local retransmission. Because it still uses the timer for lost packet recovery, it is necessary to estimate the round-trip time accurately. Besides, with a timer in a base station, a lot of system resource is wasted. At the TCP source side to prevent ACK duplication, unnecessary timeout cannot be avoided totally.

Congestion or Retransmission causes the packets to be out-of-order. If an out-of-order packet is lost at a wireless link, it is impossible to recover it with duplicated ACK or partial ACK. A current solution is the timeout retransmission, it is needed to estimate the round-trip time accurately in a wireless link, but it is difficult to do for an unreliable wireless network.

Another solution is to use a pure link layer protocol, such as automatic retransmission request (ARQ). In this solution, the protocol can give every packet a sequence number to identify the arrival order and to use it for recovery of losses, but the overhead will cause lower efficiency.

The purpose of this invention is to design a local retransmission scheme by using TCP in an unreliable network to solve the loss of out-of-order packets, by differentiating congestion loss and non-congestion loss to avoid the false actions of a TCP source. With all these effort to get a better TCP performance in an unreliable link.

Summary of the Invention:

This invention uses a technical scheme with cooperation of the transport and link layer. That is TCP Aware Local Retransmission. It uses link layer transmission sequence together with TCP source delivery sequence to detect the lost data packet, then using
5 Explicit Retransmission Notification (ERN) to avoid the false actions of a TCP source.
The invention is based on the TCP aware local retransmission.

The invention is implemented as follows: a local retransmission mechanism of TCP used in unreliable network, is characterized in the following steps:

A. When the access point of an unreliable link receives a new TCP data packet
10 from an Internet TCP source, it inserts in the packet a LAC-PDU head with a time-stamp of a first local sequence number. The packet is encapsulated to a LAC-PDU packet as "LAC-PDU head + IP head + TCP head + Data", then is delivered to the current terminal;

B. When the current terminal successfully receives a TCP data packet, it produces an acknowledgement packet (ACK1) which includes an acknowledgement number (AN).
15 It is also inserted a LAC-PDU head with the time-stamp of a second local sequence number, i.e. the acknowledgement packet is encapsulated to a LAC-PDU acknowledgement packet and is delivered back to the access point of an unreliable link;

C. At the access point of the unreliable link, a detection of whether there is a data packet loss is made. This is according to the Acknowledgement Number (AN), the Time-
20 stamp of the second local sequence number, both received from the acknowledgement package, and the Time-stamp of the first local sequence number, which is stored in the access point. If a lost data packet is detected, as step A procedure, updates its time-stamp of the first local sequence number in LAC-PDU head, and retransmits. When congestion

loss of the data packet is impossible, which corresponds to the acknowledgement number (AN) of acknowledgement packet, the acknowledgment packet (ACK1), with explicit retransmission feedback (ERN) field marked, is delivered to the TCP source.

In step A mentioned above, when a LAC-PDU head with the time-stamp of the first local sequence number is inserted, a copy of the encapsulated LAC-PDU data packet is stored in a buffer at the same time.

The time-stamp of the first local sequence number, mentioned above, is a fixed length bit field. Along with delivered data packet increases, starting from 0 with 1 as step length its value increases sequentially.

10 During the whole delivery process from access point to terminal, the real delivery sequence is uniquely determined by the time-stamp value of the first local sequence number in the TCP data packet.

The time-stamp of the second local sequence number, mentioned above, is also a fixed length bit field. It records the maximum value of the time-stamp of the first local sequence number among all the successfully received TCP data packets in the current terminal.

As mentioned above at the access point of the unreliable link, a lost data packet is detected according to the acknowledgement number (AN), the time-stamp of the second local sequence number, both comes from received acknowledgement packet, and the time-stamp of first local sequence number, which is stored at the access point. Further more, detection is made for whether the data packet is still in the access point of unreliable link, which corresponds to the acknowledgement number (AN) of acknowledgement packet, if it is, a comparison between two time-stamps is made. The

comparison is between the time-stamp of the first local sequence number in the data packet and the time-stamp of the second local sequence number in the acknowledgement packet. If the time-stamp of first local sequence number is less than the time-stamp of second local sequence number, the lost data packet is detected; then the time-stamp of first local sequence number at the LAC-PDU head is updated and the data packet is retransmitted. In addition, in the access point of the unreliable link, the data packets, which time-stamp of first local sequence number is less than the acknowledgement number (AN), are all deleted.

The update of the time-stamp of first local sequence number at the LAC-PDU head of the data packet, mentioned above, substitutes the time-stamp of first local sequence number with the current delivery sequence, then retransmits.

As mentioned above, in the access point of the unreliable link, the data packets, which TCP sequence number is less than the acknowledgement number, are all deleted at the following situations. They are:

- a data packet, which corresponds to acknowledgement number (AN) of acknowledgement packet, is not in the access point of the unreliable link;
- the time-stamp of first local sequence number is equal or greater than time-stamp of second local sequence number;
- after data packet is lost, updated and retransmitted.

The explicit retransmission (ERN) feedback, mentioned above, is a one bit field. When the data packet, corresponding to the acknowledgement number (AN), is in the access point of unreliable link, the explicit retransmission (ERN) feedback bit of acknowledgement packet (ACK1), which is to be sent to TCP source, is set. When the

TCP source receives an acknowledgement packet (ACK1) with explicit retransmission (ERN) feedback bit set, at the same time fast retransmission or timeout retransmission of the TCP data packet happen, which corresponds to the acknowledgement packet (ACK1), only the data packet is retransmitted without any shrink operation of the sending window.

5 The mentioned above time-stamp length, for first local sequence number or second local sequence number, is the maximum packet number can be buffered at the access point of unreliable link.

 The mentioned above time-stamp length, for first local sequence or second local sequence number, is an eight bit field, including one carry bit for overflow.

10 In mentioned step A, in the LAC-PDU data packet the time-stamp of first local sequence number may be substituted by a lower layer transfer sequence number, and a corresponding relationship between the sequence number of TCP data packet and its lower layer transfer sequence number is created. With this substitution, in mentioned step B, the time-stamp of second local sequence number is the successfully received
15 maximum transfer sequence number of lower layer in the terminal.

 A TCP local retransmission scheme, used in a unreliable network, the characteristic is as follows: Method of transmitting a data packet for Mobile communication system which receives the data packet including sequence number from Internet Networks and transmits the data packet to a mobile terminal; said transmitting
20 method comprising the steps of:

 receiving a new data packet from the Internet Network;

 giving a time-stamp of first local sequence number to the each received data packet;

forming the packet data to a format used in the mobile communication system
with the given time-stamp of first local sequence number;

buffering the packet data with the time-stamp;

transmitting the formed packet data to the mobile terminal sending back an

- 5 acknowledgement data from the mobile terminal, when the mobile terminal successfully
receives the transmitted packet data, which includes an acknowledgement number and the
time-stamp of second local sequence number corresponding to the received time-stamp of
first local sequence number;

detecting the data packet which should be retransmitted to the mobile terminal by

- 10 comparing the buffered sequence number and the time-stamp of first sequence number
with the sent back acknowledgement data; and

retransmitting the detected data packet.

Method of transmitting a data packet, wherein said mobile communication system
includes a server, said server gives the time-stamp of first local sequence number to each
15 received data packet, forms the packet data to the format used in the mobile
communication system with the time-stamp of first local sequence number and buffers
the packet data with the time stamp.

A TCP local retransmission scheme, used in an unreliable network, the
characteristic is as follows: Method of receiving a data packet of a Mobile terminal in a
20 Mobile communication system which receives the data packet including a sequence
number from Internet Networks and transmits the data packet to the mobile terminal; said
receiving method comprising the steps of: receiving a new data packet which is formed
to a format used in the mobile communication system, the formed data packet includes a

time-stamp of first local sequence number; giving a time-stamp of second local sequence number corresponding to the received time-stamp of first local sequence number;

5 sending back an acknowledgement data which includes an acknowledgement number and the given time-stamp of second local sequence number corresponding to the received time-stamp of first local sequence number.

The invention points to the user terminal environment, such as a mobile terminal. With an unreliable link, such as a wireless link, it is connected to a data network, such as the Internet, and a one direction TCP connection from remote server to wireless access point then to mobile terminal is established. The invention solves the non-congestion loss

10 problem of the out-of-order TCP data packet in the unreliable link with the following ideas. The time-stamp of first local sequence number is used to identify the real transmitting sequence of the TCP data packets in the unreliable link segment. Then with the acknowledgement number (AN) of the TCP acknowledgement packet and the corresponding first and second local sequence number, the lost data packet, which
15 corresponds to the acknowledgement number (AN), is detected. If it is lost, the data packet is retransmitted locally. At the same time, with explicit retransmission (ERN) feedback, false window action is prevented. Therefore, the invention is one kind of cooperation, between time-stamp of local sequence number and acknowledgement number (AN), to detect the lost data packet, and it is also an explicit retransmission
20 feedback method.

The advantages of this invention are as following. Because the lost data packet is detected by the cooperation of the TCP acknowledgement packet (ACK1) and local sequence number time-stamp, and it is retransmitted locally, so without the lower layer

extra overhead for control, the non-congestion loss of data packet is recovery by retransmission in time. It has characteristics of fast recovery and high efficiency. It recovers out-of-order TCP lost data packet effectively.

5 Brief Description of the Drawings:

Figure 1 is the structure diagram of one direction TCP connection in a mixed network with wireless access;

Figure 2 is the structure diagram of data packet transmission from wireless access point to mobile terminal;

10 Figure 3 is the corresponding relationship diagram between TCP sequence number and time-stamp of first local sequence number; and

Figure 4 is the flowchart diagram using acknowledgement packet (ACK1) to detect data packet loss.

15 Detailed Description:

Reference to figure 1, a mixed network with wireless access, for example mobile communication system based on IMT2000, is shown. 10 is the international interconnection network (Internet Network), 20 is a remote server, i.e. TCP source, of Internet, 30 is a wireless (unreliable link) access point server, 40 is a mobile terminal (also a mobile computer, user terminal or user computer). The wireless access server is in
20 a mobile communication system, which comprises a Mobile Exchanger, Radio Network Controller and base station. The wireless access server can be set up in the Mobile Exchanger, Radio Network Controller or base station. Further, functions of the wireless

access server can be installed to the Mobile Exchanger, Radio Network Controller or base station without setting up the wireless access server.

The operation of the wireless access point server 30 is as following. When it receives a new TCP data packet which comes from TCP source 20 in Internet 10, it encapsulates the data packet in a Link Access Control-Protocol Data Units (LAC-PDU). There is a fixed length bit field in the LAC-PDU head for each LAC-PDU data packet. This field is used to record the real transmission sequence number, i.e. delivery sequence, and is referred to as time-stamp of first local sequence number. The structure of the LAC-PDU data packet is shown in figure 2. It is "LAC-PDU head (with time-stamp of first local sequence number) + IP head + TCP head + DATA". When the time-stamp of first local sequence number is filled in, the LAC-PDU data packet is also copied to the buffer at the same time.

As data packets increase, the time-stamp value of first local sequence number is increased with the step of 1, starting from 0. Therefore each TCP data packet moves from wireless access point server (base station) 30 to mobile terminal 40, its real transmitting sequence is uniquely defined by the time-stamp of first local sequence number in LAC-PDU data packet. As shown in figure 3, #1, #2, #4, #3 ... represent a TCP data packet sequence number; 0, 1, 2, 3 ... represent time-stamp value of first local sequence number corresponding to LAC-PDU data packet #1, #2, #4, #3 ... respectively. They are delivered along the time axis t sequentially. In the international interconnection (i.e. Internet) or large scale local area network (i.e. WAN: wide area network), there are many paths to connect between a data source and end user to realize a reliable network. This means that if the first established path has network congestion, the system may choose

another path to deliver data. However, each path has a different delivery time. Thus, during a data transfer, a path change may cause out-of-order data delivery. As a result, there is a case that the wireless access server receives TCP packet data by above order of #1, #2, #4, #3

5 Mobile terminal 40 produces an acknowledgement packet (ACK1) for each successfully received TCP data packet. As the TCP data packet, the acknowledgement packet is also encapsulated in a LAC-PDU. At the head of this LAC-PDU packet, there is a field for time-stamp of second local sequence number. This field records the maximum time-stamp value of first local sequence number for all successfully received TCP data
10 packet. The current mobile terminal sends back to wireless access point server (base station) 30 acknowledgement packet (ACK1), with the time-stamp of second local sequence number and acknowledgement number. It is used to detect TCP data packet loss at the wireless link. The detection method is further described by reference to figure 4. Said Acknowledgment number means a number that indicates the next sequence number
15 expected. In a backhanded manner, this also shows the TCP sequence number of the continuously received last TCP data packet; it shows TCP sequence number of the continuously received last TCP packet data plus 1.

Step 401 detects whether there is a TCP data packet, which sequence number corresponds to the acknowledgement number (AN) of acknowledgement packet (ACK1),
20 in the buffer of wireless access point server 30 (base station). If it is, then executes step 402, otherwise step 406. Step 402 gets the time-stamp of first local sequence number in the TCP data packet, said ts. Step 403 compares ts with the time-stamp of second local sequence number, said ts_{max} , carried by the acknowledgement packet, to detect whether

there is $ts < tS_{max}$. If it is $ts < tS_{max}$, then executes step 404, 405 and 406, otherwise step 406. When $ts < tS_{max}$, step 404 and 405 are executed. It means the TCP data packet, corresponding to acknowledged data packet sequence number (AN), is lost. Then substituting the time-stamp of first local sequence number, the TCP data packet is

5 retransmitted. This means a new time-stamp, which represents current transmitting sequence number, substitutes the time-stamp of first local sequence number of the LAC-PDU data packet, which corresponds to the acknowledgement number (AN), then the updated LAC-PDU data packet is retransmitted. Step 406 deletes all the TCP data packet, which sequence number is less than AN, in the wireless access point server (base station)

10 30, under the following situations. They are: there isn't a TCP data packet, corresponding to acknowledged data packet sequence number (AN) in acknowledgement packet (ACK1), in the buffer of wireless access point server (base station) 30; there isn't $ts < tS_{max}$ and after step 404 and 405 are executed.

In the invention, an explicit retransmission (ERN) feedback (or explicit

15 retransmission notification) bit is defined in the unused bits (as 6 bits) of acknowledgement packet (ACK1) head. When there is a TCP data packet, corresponding to the acknowledged data packet sequence number in the acknowledgement packet, in the buffer of wireless access point server (base station) 30, the explicit retransmission (ERN) feedback bit in this acknowledgement packet is set. The acknowledgement packet

20 (ACK1), which includes the explicit retransmission (ERN) feedback and the acknowledged data packet sequence number AN, is sent to TCP source 20. When TCP source 20 receives an acknowledgement packet with explicit retransmission (ERN) feedback bit set, and the TCP data packet, corresponding to the acknowledgement packet

(ACK1), is in the fast retransmission or timeout retransmission, there are no operations to shrink the window size, only the TCP data packet is retransmitted.

The time-stamp value of first or second local sequence number is fixed. Its length is the possible maximum buffered data packets for a TCP connection in the wireless access point (base station). If a TCP data packet length is 576 bytes, time-stamp of local sequence number takes value 0 - 127 (in general, it is enough for a TCP connection), i.e. time-stamp length of sequence number is 7 bits.

As the transmitted data packet number increases, when the time-stamp value of local sequence number exceeds its possible represented maximum value, an overflow happens and it is counted starting from 0 again. In this case, the time-stamp of local sequence number cannot represent the real data packet sequence number, so there is a carry bit in this invention. When overflow happens, the carry bit is complemented. In this way at step 403, a correct checking for greater and less relation of two time-stamps of local sequence number can be made, so an 8 bits length of time-stamp for local sequence number is suggested.

In the invention, a time-stamp of local sequence number is used to identify real sequence number of TCP data packet transmitted on an unreliable link, so any other equivalent manners can be employed, if it implements the function.

The implementation of this invention encapsulates a TCP data packet and time-stamp of local sequence number together in LAC-PDU, then transmits. In many mobile communication systems, they provide transfer function of transmitting sequence number in lower layer (data link layer). In this case the detection of data packet loss can be performed under the following conditions. They are: corresponding relationship between

sequence number of TCP data packet and transmitting sequence number of its lower layer is established; the mobile terminal sends back to the base station the lower layer maximum transmitting sequence number it receives. In practice, the lower layer transmitting sequence number acts as the function of time-stamp of local sequence number.

The invention can be widely used in many reliable data communication which access data network - Internet with an unreliable link.

The key idea of this technical scheme is combining the advantages of TCP and Data Link layer, in order to decrease the data link layer overhead and provide effective reliable transmission for out-of-order packet. In this scheme, the time-stamp, which can be seen as sequence number of data link layer for the TCP section, represents the TCP section transmitting sequence in the unreliable link access point.